Three-Dimensional Fitt's Law Model Used to Predict Movement Time in Serious Games for Rehabilitation

Sergio García-Vergara and Ayanna M. Howard

Georgia Institute of Technology, Electrical and Computer Engineering Department, 85 5th Street NW, Atlanta, GA 30332
sergio.garcia@gatech.edu, ayanna.howard@ece.gatech.edu

Abstract. Virtual reality serious game platforms have been developed to enhance the effectiveness of rehabilitation protocols for those with motor skill disorders. Such systems increase the user's motivation to perform the recommended in-home therapy exercises, but typically don't incorporate an objective method for assessing the user's outcome metrics. We expand on the commonly used human modeling method, Fitt's law, used to predict the amount of time needed to complete a task, and apply it as an assessment method for virtual environments. During game-play, we compare the user's movement time to the predicted value as a means for assessing the individual's kinematic performance. Taking into consideration the structure of virtual gaming environments, we expand the nominal Fitt's model to one that makes accurate time predictions for three-dimensional movements. Results show that the three-dimensional refinement made to the Fitt's model makes better predictions when interacting with virtual gaming platforms than its two-dimensional counterpart.

Keywords: Fitt's law, virtual reality games, physical therapy and rehabilitation, linear modeling.

1 Introduction

Gaming platforms for serious games play an important role in the rehabilitation field [1]. Such systems have been developed to increase the motivation of users to perform their in-home recommended exercises [2], [3]. Moreover, previous research has shown these systems can be used to calculate kinematic metrics associated with an individual's movement profile. In [4], a prototype rehabilitation game was presented that used the Kinect system to analyze biomechanical movements of the upper extremities represented as range of motion and posture data. In [5], an augmented reality system that enabled 3D-reaching movements within the environment was presented. They derived a set of kinematic data represented as movement time and end-effector curvature values. Finally, [6] evaluated the probability of recognizing six different movement gestures, useful for rehabilitation, when using a virtual gaming system. Although virtual systems such as these show the viability of collecting kinematic movement data, they do not provide a quantifiable means of determining the quality of that movement. As such, we focus on incorporating a methodology within existing virtual reality (VR) gaming platforms that objectively evaluates the outcome metrics of an individual during game play.

A common symptom experienced by individuals who have a motor skill disorder is slow movements [7]. As such, movement time (MT) – defined as the time needed to complete a given task – is a kinematic parameter of interest in rehabilitation interventions because it directly correlates with the speed of the individual's movements. In this paper, we focus on predicting the MT needed to complete a task in any VR gaming platform. We use the prediction as the ground truth value for quantitatively comparing the user's nominal MT as a means for assessing their kinematic performance. Because of its wide adoption, we make use of the model of human movement, Fitt's law [8]. This law predicts the amount of time a user needs to reach a given target in a virtual environment. Even though refinements to improve the accuracy of Fitt's law have been made to the original model, to the best of our knowledge, there has not been any directly derived for time prediction for three-dimensional (3D) movements; which are inevitable when interacting with a VR system or a serious gaming platform.

As such, we propose a new variation on the Fitt's law model that takes into consideration 3D spatial movements. Section 2 presents a short literary review on previous variations and modifications made to the original model. Section 3 discusses in detail the procedure taken to create our final model. Section 4 presents the results obtained in testing sessions with human participants. Finally, we analyze the results in Section 5, and make our concluding remarks in Section 6.

2 Background

Fitt's law was initially designed to predict the amount of time a user needs to complete a task in order to design better human-computer interaction (HCI) interfaces or to determine the best input method for a digital system. Card et al. [9] used Fitt's law to evaluate four devices with respect to how rapidly they can be used to select text on a CRT display.

Walker et al. [10] compared selection times between walking menus and pull-down menus. Gillian et al. [11] used Fitt's law to examine the needed time to select a text using a movement sequence of pointing and dragging.

Even in these applications for Fitt's law, HCI researchers have developed several refinements to improve the accuracy of the model. MacKenzie [12] summarized some refinements that deal with the definition of the difficulty of a task. In the original model, the difficulty of a movement task (*DI* for "difficulty index"), was quantified by (1).

$$DI = \log_2(2 * A/W) \tag{1}$$

where A is the distance to move, and W is the width of the target to reach. Welford [13] proposed a new formulation for DI (2) after noting a consistent departure of data points above the regression line for 'easy' tasks (i.e. DI < 3 bits).

$$DI = \log_2(A/W + 0.5) \tag{2}$$

Moreover, a preferred formulation (3), known as the Shannon formulation [14], is commonly used because it provides a better fit with observations, mimics the information theorem underlying Fitt's law, and provides a positive rating for the DI.

$$DI = \log_2(A/W + 1) \tag{3}$$

To the best of our knowledge, none of the previous studies have used Fitt's law for human movement assessment purposes, and the supporting literature for the theory behind Fitt's law is limited to two-dimensional (2D) movements. In this paper we discuss a methodology for building a model that: 1) predicts movement time for three-dimensional movements, and 2) is used as a tool for quantitatively assessing an individual's kinematic performance.

3 Methodology

3.1 Serious Game Platforms

In this paper we focus on expanding the functionality of serious game platforms used for rehabilitation by incorporating an objective kinematic assessment methodology. We make use of the developed platform called

Super Pop VR^{TM} [15], [16]. It combines interactive game play for evoking user movement with an objective and quantifiable kinematic algorithm to analyze the user's upper-arm movements in real-time. While engaged with the game, users are asked to move their arms to 'pop' virtual bubbles of various sizes, which appear at various locations in the virtual environment. As the bubbles appear on screen, a 3D depth camera maps the user's movements into the virtual environment. These movements map into movement tasks that require reaching a target from a specified initial position; which are evaluated by Fitt's law. Figure 1 shows a comparison between a reaching task evaluated by Fitt's law (Figure 1a), and an example of a reaching exercise in the *Super Pop VR^{TM}* platform (Figure 1b). The ability to reach is critical for most, if not all, activities of daily living such as feeding, grooming, and dressing [17]. Failure to recover upper-extremity function can lead to depression [18]. As such, reaching movements, correlated to reaching exercises, are of interest in various rehabilitation scenarios.

Applying Fitt's law to the Super Pop game, we focus on predicting the amount of time a user needs to move between two displayed 'bubbles' as a function of the distance between the virtual objects and the width (diameter) of the target 'bubble'. Given the nature of the described platform, users move their arms in the 3D space in order to interact with the virtual objects on the screen. As such, we first need to build an appropriate model (i.e. define a DI), that is able to make accurate time predictions for 3D movements.

(a) (b)

Fig. 1. Comparison between a common movement task evaluated by Fitt's law (where A is the distance traveled, and W is the width of the target) (a), and a reaching exercise in the Super Pop VR^{TM} platform (b). Figure (a) adapted from [12].

3.2 Linear Models

Fitt's law predicts movement time as a linear function of the difficulty index (DI) of a task (4). Because of its wide adoption and popularity, we adhere to the DI definition of the Shannon formulation as seen in (3), resulting in a model of the form (5).

$$MT = a + b * DI \tag{4}$$

$$MT = a + b * \log_2(A/W + 1) \tag{5}$$

where MT is the predicted movement time (in milliseconds), A is the distance to move, W is the width of the target to reach, and a and b are the intercept and the slope of the model respectively. Building a Fitt's model refers to training the slope and intercept to fit MT data collected from users interacting with the system. In general, a number of movement tasks are defined by selecting different combinations of traveled distances and widths of targets, and then by calculating the corresponding DI. Human MT data are collected for each defined task, and a linear regression between the MT averages per task and their corresponding DIs is performed to compute the slope and intercept of the model.

Since we are interested in building a model that is appropriate for 3D movements, the distance travelled is now the 3D Euclidean distance between the initial position of the user's hand, and the target. However, a complication arises because the positions of the 'bubbles' in the virtual platforms are defined in a 2D space. This means that the movement tasks are selected based on 2D data. As such, we built two linear models. The first model correlates the 2D pixel distance between the virtual objects to the user's 3D path length (PL). We then use this model to calculate the distance travelled parameter in (5) and create our second model: the correlation between the DI of a movement task and the time needed to complete it.

For collecting human MT data needed to develop the model, we recruited seven able-bodied adults to interact with the *Super Pop VR*^{TM} game. Sixteen tasks were empirically selected; each participant was assigned to repeatedly complete eight of them. We collected, on average, 24 ± 5 PL and MT points for each task. To increase the correlation factor between variables for both models, we assume that both datasets follow a Gaussian distribution, and thus only considered data points that were within one standard deviation of the mean of the complete dataset. Moreover, taking into consideration the learning curve of the platform, the

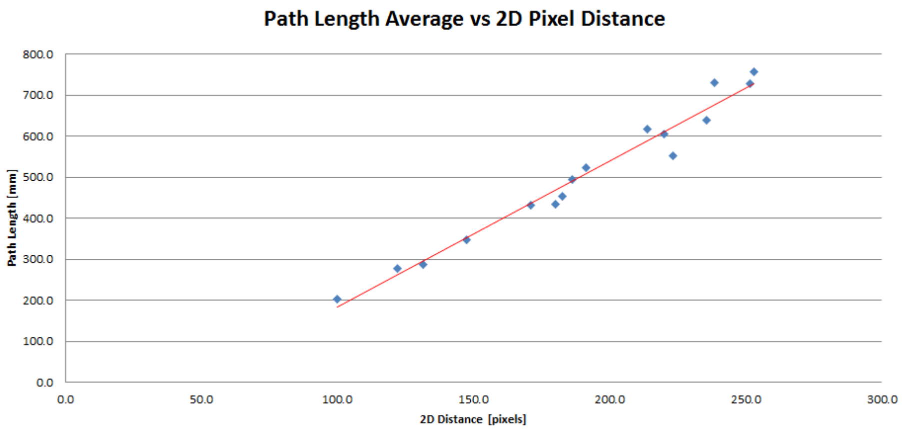

Fig. 2. 3D path length averages of collected human data versus 2D pixel distance between virtual objects. Figure also shows the final linear correlation (continuous line) between the two variables.

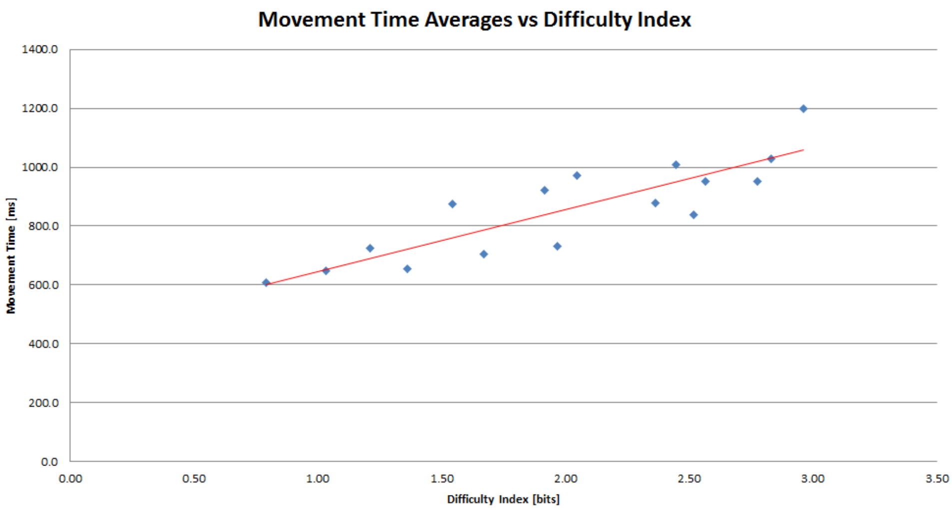

Fig. 3. Movement time averages collected from human data versus the corresponding DIs for each task. Figure also shows the final linear correlation (continuous line) between the two variables.

participants were required to interact with the system twice before starting the actual collection of the data. This practice eliminates the possible errors due to unfamiliarity with the game.

A linear regression was performed on the collected path length data to correlate the participants' 3D PL to the selected 3D pixel distance between 'bubbles' (Figure 2); which yielded (6) with a correlation factor of $\mathbf{R^2=0.9703}$. MacKenzie [12] argues that correlations above 0.900 are considered to be very high for any experiment involving measurements on human subjects. Thus, we can conclude that the PL model provides a good description of the observed behavior.

$$PL = 3.5651 * D_p - 174.3 \tag{6}$$

where D_p is the 2D distance between the two virtual targets, and PL is the 3D path length travelled by the user for the corresponding D_p in mm.

A second linear regression was performed on the collected MT data to correlate the participants' MT to the DI of the corresponding tasks (Figure 3); which yielded (7) with a correlation of $\mathbf{R^2=0.7428}$. Although the resulting correlation factor is not considered to be 'very high', it still suggests that the MT model also provides a good description of the observed behavior. The DI of the tasks was calculating using (5), making use of the built PL model (6) to substitute for the travelled distance.

$$MT = 208.97 * DI + 435.02 \tag{7}$$

where DI is the difficulty index of a given task, and MT is the movement time prediction made for the task (in milliseconds). It's important to mention that (7) is limited to the selected definition of DI. If a different definition were to be used, the MT model would have to be retrained.

Combining equations (5), (6), and (7), we obtain the final MT model (8) as a linear function of the 2D pixel distance between two virtual objects by making use of a second linear model of human PL data.

$$\boxed{MT = 208.97 * \log_2\left(\frac{3.5651 * D_p - 174.3}{W} + 1\right) + 435.02} \tag{8}$$

where D_p is the 2D pixel distance between two virtual objects of the given task, W is the width of the second virtual object, and MT is the movement time prediction made for the given task (in milliseconds). Since the argument of the logarithm has to be unit less and since the PL model computes values in millimeters, the width of the target has to also be in millimeters.

In order to better determine the accuracy of the 3D Fitt's model, we also created a common 2D Fitt's model and compared the prediction results to the nominal MT values collected from the participants. The 2D model was built in a similar fashion than the 3D model. The selected tasks were the same as those for the 3D model. The difference relies on the fact that the DIs for the tasks were computed using the 2D pixel distance directly, instead of applying the PL model. A linear regression was applied to the collected MT data to correlate the participants' MTs to the DIs of the corresponding tasks; which yielded (9) with a correlation factor of $\mathbf{R^2=0.7346}$.

$$MT = 245.2 * DI + 377.42 \tag{9}$$

4 Experimental Results

The final model was tested with seventeen able-bodied high school students. Seven females and ten males ranging in age between 15 and 16 years (mean age = 15.5 years, standard deviation = 0.5 years) were recruited to interact with the *Super Pop VR^{TM}* game in order to validate that the proposed methodology for creating a Fitt's law model is appropriate for 3D movements. The participants interacted in an office setting, which was maintained constant in order to maintain consistency. The virtual reality game screen was projected onto a large screen via a projector connected to a PC laptop. The chair height upon which the participants sat was 41cm tall, the distance between the user's chair and the depth camera was 190cm, and the distance between the projector and the screen was 170cm. Each participant was asked to play four games (two per arm), and PL and MT was collected for a total of six trials per arm.

Taking into consideration the learning curve of the used platform, we evaluate the last trial of the participants' dominant hand. Table 1 summarizes a comparison between the participants' nominal MT for the selected trial and the movement time predictions made by the 2D and 3D models. The error of the prediction is defined as the absolute difference between the participant's MT and the prediction made. The table also shows which model made the best prediction for each case; the model that best fits the given scenario is the model with the smallest difference between the actual MT and prediction. Figure 4 expands on Table 1 by organizing the results in a graphical medium.

Table 2 shows the progression of MT values over the six trials of Participant 2's dominant hand. Similar to Table 1, Table 2 shows a comparison between Participant 2's nominal MT and the predictions made by the 2D and 3D models. The table also includes the decision of the model that makes the most accurate prediction based on the absolute difference between the actual MT value and the predictions made.

Table 3 shows a summary of how the models behave for clear 2D and 3D movements. The data collected from the last trial of Participant 7's dominant hand are considered as 3D movements, while the data collected from the last trial of Participant 16's dominant hand are considered as 2D movements. The table shows the MT predictions made by both models on the two described scenarios, and the participants' actual MT for both scenarios.

5 Analysis

It's important to keep in mind that the linear models were built with data collected from adults. This allows for the possibility of over (or under) predicting path length (PL) and movement time (MT) values given that they were tested with data collected from high-school teenagers. Previous research has been shown that kinematic capabilities, among other parameters, are a nonlinear function of the age of the individual [19]. As such, there are some scenarios where neither the 2D nor 3D models make accurate MT predictions. For example, Participant 5 moved in almost double the time than what both models predicted (Table 1). Moreover, we only collected 24 ± 5 data points per task, while studies similar to [20] collected 470 trials per task. More data would results in a higher correlation and, thus, more stable models.

Another important observation is that our 3D Fitt's model falls into the known two-dimensional model when the movements are (almost) planar. Table 1 shows that both models make very similar predictions in these scenarios, suggesting that there is no deterioration when applying the 3D model to 2D movements. More importantly, in scenarios where the individual makes 3D movements, the 3D model makes more accurate predictions than the 2D model. Table 3 shows an example of such scenarios. The table shows that the predictions made by both models for a case where the movements were in the 2D space (participant 16), are relatively similar to each other. The difference between the predictions is **5.22 ms**, the difference between the prediction of the 3D model and the actual MT is **208.28 ms**, and the difference between the prediction of the 2D model and the actual MT is **213.51 ms**.

Similarly, Table 3 shows that the prediction made by the 3D model is more accurate than the prediction made by the 2D model for a case where the movements where in the 3D space. The difference between the two predictions is **411.41 ms** (which is considerably of greater value than that of the 2D movements), the difference between the prediction made by the 3D model and the actual MT is **35.96 ms**, and the difference between the prediction made by the 2D model and the actual MT is **375.45 ms**. These results show that, for 3D movements, the proposed Fitt's model with a PL model included makes more accurate MT predictions than the original Fitt's model.

Table 1. Comparison between the participants' MT nominal values and the predictions made by the 2D and 3D models to determine the best predictor for each scenario.

Participants	MT_U [ms]	2D DI [bits]	2D MT_p [ms]	Difference [ms]	3D DI [bits]	3D MT_p [ms]	Difference [ms]	Best Predictor
1	1129.69	1.44	731.67	398.02	1.33	713.42	416.27	2
2	901.69	1.46	735.63	166.05	1.54	757.77	143.92	3
3	1180.66	1.01	624.60	556.07	2.73	1005.02	175.64	3
4	980.43	1.27	689.89	290.54	1.48	743.60	236.83	3
5	1413.22	1.68	789.09	624.13	1.48	745.13	668.09	2
6	1283.64	1.54	754.32	529.32	1.60	770.16	513.48	3
7	958.87	0.84	583.42	375.45	2.68	994.83	35.96	3
8	1071.39	1.46	734.25	337.14	1.99	850.30	221.09	3
9	978.12	1.17	665.26	312.85	1.10	664.87	313.25	2
10	913.74	1.80	819.61	94.13	1.64	777.94	135.80	2
11	1167.51	1.50	745.69	421.83	1.34	715.61	451.91	2
12	943.35	1.50	745.43	197.92	1.81	812.25	131.10	3
13	926.22	1.86	834.28	91.93	1.76	802.81	123.41	2
14	949.05	1.31	699.43	249.62	2.40	935.79	13.26	3
15	1352.22	1.59	767.88	584.34	1.56	760.11	592.11	2
16	934.13	1.40	720.62	213.51	1.39	725.84	208.28	3
17	1046.67	1.55	756.61	290.05	1.46	739.80	306.86	2

Table 2. Progression of MT values over the six trials of participant 2's dominant hand.

Trials	MT_U [ms]	2D DI [bits]	2D MT_p [ms]	Difference [ms]	3D DI [bits]	3D MT_p [ms]	Difference [ms]	Difference [ms]	Best Predictor
1	1055.81	1.38	716.67	339.13	1.59	768.14	287.66	51.47	3
2	1352.83	1.31	697.57	655.26	1.32	711.19	641.64	13.62	3
3	1010.50	1.76	808.77	201.73	1.89	830.98	179.52	22.21	3
4	972.78	1.51	748.32	224.45	1.50	748.91	223.87	0.59	3
5	898.18	1.43	727.28	170.89	1.87	824.94	73.23	97.66	3
6	901.69	1.46	735.63	166.05	1.54	757.77	143.92	22.14	3

Table 3. Predictions made by both models when applied to clear 2D and 3D movements.

	2D Movement (Participant 16)	3D Movement (Participant 7)
Prediction from 2D Model [ms]	720.62	583.42
Prediction from 3D Model [ms]	725.84	994.83
Participant's Actual MT [ms]	934.13	958.87
Difference between predictions [ms]	5.22	411.41
Difference between 2D prediction and actual MT [ms]	213.51	375.45
Difference between 3D prediction and actual MT [ms]	208.28	35.96

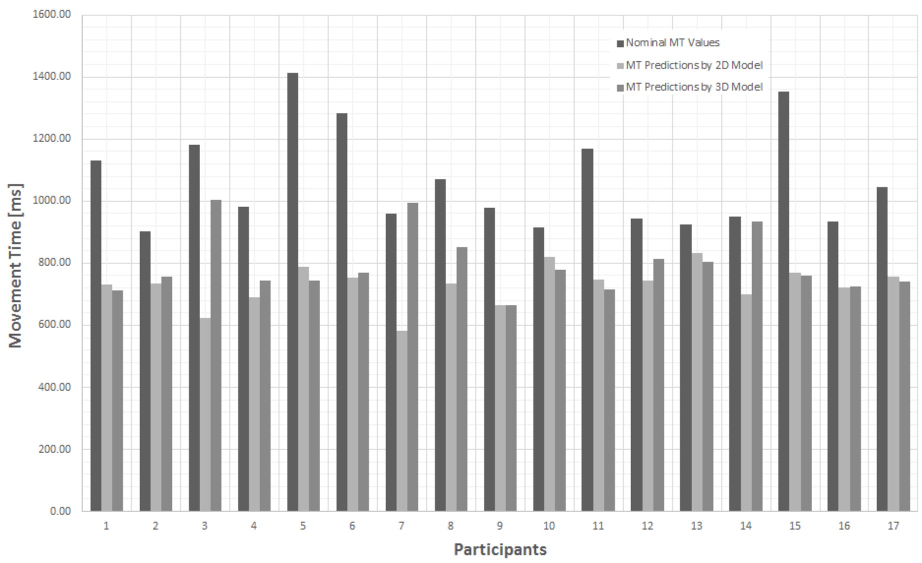

Fig. 4. Comparison between the participants' nominal MT and the predictions made by the 2D and 3D models.

6 Conclusion and Future Work

The proposed methodology for developing a 3D Fitt's law model has the potential to be incorporated into existing serious game platforms as an effective means of rehabilitation. Results show that the final 3D model can better predict human MT for different reaching tasks when compared to its 2D counterpart. For future consideration, in order to have a fully robust methodology for 3D time prediction model, data from different age demographics must also be collected and added to the model.

Acknowledgements. This work was supported in part by the NSF Graduate Research Fellowship under Grant No. DGE-1148903, and NSF Grant 1208287. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author's and do not necessarily reflect the views of the National Science Foundation.

References

1. S. Taria, M. Johanneson, and P. Backlund, "Serious Games: An overview," 2007.
2. K. Harris, and D. Reid, "The influence of virtual reality on children's motivation," *Canadian Journal of Occupational Therapy*, Vol. 72, No. 1, pp. 21-29, 2005.

3. D. Reid, "The influence of virtual reality on playfulness in children with cerebral palsy: A pilot study," *Occupational Therapy International*, Vol. 11, No. 3, pp. 131-144, 2004.
4. D. Q. Freitas, A. Da Gama, L. Figueiredo, T. M. Chaves, D. Marques-Oliveira, V. Teichrieb, and C. Araújo, "Development and Evaluation of a Kinect Based Motor Rehabilitation Game," *Brazilian Symposium on Computer. Games and Digital Entertainment (SBGames 2012)*, Brazil, pp. 144-153, 2012.
5. M. Goffredo, M. Schmid, S. Conforto, and T. D'Alessio, "3D reaching in Visual Augmented Reality using Kinect™: the Perception of Virtual Target," *Proc. International Conference on NeuroRehabilitation ICNR*, Vol. 2, pp. 711-715, 2012.
6. E. Davaasambuu, C. C. Chiang, J. Y. Chiang, Y. F. Chen, and S. Bilgee, "A Microsoft Kinect based virtual rehabilitation system," *in Proceedings of the 5th International Conference (FITAT '12),* pp. 44-50, 2012.
7. M. E. Morris, "Movement disorders in people with Parkinson Disease: A model for physical therapy," *Physical Therapy*, Vol. 80, No. 6, pp. 578-597, 2000.
8. P. M. Fitts, "Perceptual-motor skill learning," *Categories of Human Learning*, Vol. 47, pp. 381-391, 1964.
9. S. K. Card, W. K. English, and B. J. Burr, "Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT," *Ergonomics*, Vol. 21, No. 8, pp. 601-613, 1978.
10. N. Walker, and J. B. Smelcer, "A comparison of selection times from walking to pull-down menus," *Proceedings of the CHI Conference on Human Factors in Computing Science*, pp. 221-225, 1990.
11. D. J. Gillian, K. Holden, S. Adam, M. Rudisill, and L. Magee, "How does Fitt's Law fit pointing and dragging?," *in Proceedins of the CHI Conference on Human Factors in Computing Systems*, pp. 227-234, 1990.
12. I. S. MacKenzie, "Movement time prediction in human-computer interfaces: A brief tour on Fitt's Law," *Proceedings Graphics Interface*, Vol. 92, pp. 140-150, 1992.
13. A. T. Welford, "The measurement of sensory-motor performance: Surbey and reappraisal of twelve years' progress," *Ergonomics*, Vol. 3, pp. 189-230, 1960.
14. C. E. Shannon, and W. Weaver, "The mathematical theory of communication," Urbana, ILL University of Illinois Press.
15. S. García-Vergara, Y. P. Chen, and A. M. Howard, "Super Pop VR™: An adaptable virtual reality game for upper-body rehabilitation," *in Virtual, Augmented, and Mixed Reality: Systems and Applications* (Springer Berlin Heidelberg), pp. 40-49, 2013.
16. S. García-Vergara, L. Brown, H. W. Park, and. A. M. Howard, "Engaging children in play therapy: The coupling of virtual reality games with social robotics," *in Technologies of Inclusive Well-Being* (Springer Berlin Heidelberg), pp. 139-163, 2014.
17. C. V. Granger, B. B. Hamilton, and F. S. Sherwin, "Guide for the user of the uniform data set for medical rehabilitation," *Uniform data system for medical rehabilitation project, Buffalo General Hospital, New York*, 1986.
18. P. H. McCrea, J. J. Eng, and A. J. Hodgson, "Biomechanics of reaching: clinical implications for individuals with acquired brain injury," *Disability & Rehabilitation*, Vol. 24, No. 10, pp. 534-541, 2002.
19. C. J. Ketcham, and G. E. Stelmach, "Age-related declines in motor control," *Handbook of the psychology of aging*, Vol. 5, pp. 313-348, 2001.
20. I. S. MacKenzie, A. Sellen, and W. Buxton, "A comparison of input devices in elemental pointing and dragging tasks," *Proceedings of the CHI Conference of Human Factors in Computing Systems*, pp. 161-166, 1991.